DOCUMENT RESUME

ED 254 538 TM 850 049

AUTHOR Saretzky, Gary D.
TITLE Treatment of Scores of Questionable Validity: The Origins and Development of the ETS Board of Review--ETS Archives Occasional Paper.
INSTITUTION Educational Testing Service, Princeton, N.J.
PUB DATE 7 Sep 84
NOTE 21p.
PUB TYPE Historical Materials (060)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Achievement Tests; *Administrative Policy; Cheating; *College Entrance Examinations; Organizational Change; *Scores; *Testing Problems; Testing Programs; *Validity
IDENTIFIERS *Educational Testing Service; *Review Panels

ABSTRACT

This report provides historical background on the origins, development and procedures, of Educational Testing Service's (ETS's) Board of Review. Established in 1969, the Board of Review makes final decisions for all test scores of questionable validity. ETS cancels or withholds scores believed to be invalid. Reasons for invalid scores range from improper testing conditions to overt candidate misconduct. Between these extremes lie the problems of questionable validity. Cheating on tests has been a recognized problem since the College Entrance Examination Board's founding in 1901. After the 1947 establishment of ETS, first test program directors, then a security officer (1956) were responsible for test score investigations. In the 1960's the Law School Admission Council's concern led to a review of ETS test security procedures directed by Robert Smith, and the subsequent establishment of the Board of Review. While there have been ongoing policy changes and procedural refinements (most notably increasing reliance on sophisticated statistical methods and computer technology) the key element is unchanged. ETS interest is in the validity of the scores it reports, not in providing evidence or judgments of candidate misconduct. Court cases concerning score reporting have upheld ETS policies. (BS)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED254538

TREATMENT OF SCORES OF QUESTIONABLE VALIDITY: THE ORIGINS AND DEVELOPMENT OF THE ETS BOARD OF REVIEW

Gary D. Saretzky

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

ETS Archives Occasional Paper

Educational Testing Service

Princeton, New Jersey

1984

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

G. Saretzky

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

TM 850 049

Treatment of Scores of Questionable Validity:

The Origins and Development of the ETS Board of Review

Educational Testing Service is obligated to test takers, score recipients, and test sponsors to report scores which reasonably quantify the abilities of of examinees. Concomitantly, ETS cancels or withholds scores it believes to be invalid and, no less if not more significantly, establishes safeguards designed to prevent erroneous cancellations of valid scores. These obligations derive from ETS's status as a not-for-profit educational organization created to serve the public interest; from the ethical standards of the professional organizations to which ETS and its staff members belong; from the contracts under which ETS performs its services; and from the implied agreements with individual candidates to report valid scores in return for their registration fees.

Invalid scores may result from defective test materials, mistimings, or other irregularities for which a test taker cannot be deemed responsible. Invalidity may also result from deliberate examinee misconduct. ETS has consistently authorized its test supervisors to dismiss candidates caught in the act of using prohibited reference materials, giving or receiving assistance, or other prohibited behavior. Similarly, ETS may cancel scores after a test administration if it receives convincing evidence of misconduct.

Between these two extremes of improper testing conditions and overt candidate misconduct lies the difficult problem of scores of questionable validity. The validity of a score may be questionable for reasons such as the following: a report that misconduct may have occurred; an improbable score gain over a previous test; an unusual agreement with responses on another test paper; striking differences in handwriting among test materials

purportedly written by the same candidate; or an indication that a test taker may have had access to test questions before the examination (preknowledge).

In 1969, the ETS Trustees adopted a new policy governing treatment of questionably valid scores. Through the establishment of the Board of Review, it substantially strengthened the organization's assurance that decisions concerning the validity of test scores would be made with the utmost care; that all test takers whose scores were investigated would be treated fairly; and that such score cancellations would be motivated solely by a concern for score validity and not by an intent to penalize candidates who may have violated rules of test administration. To put the Board's origins in perspective, it may be helpful to review earlier test security practices.

By the time the College Entrance Examination Board was founded in 1900, cheating on tests was a recognized phenomenon of human behavior.[1] That the College Board founders were aware that test takers might engage in misconduct is manifestly evidenced by the Board's Document No. 2 of February 1, 1901. The section entitled "Instructions to Candidates for Examination" warned the first College Board candidates that the presence of contraband material would be cause for dismissal from the examination room, as would be giving or receiving assistance. "Upon this subject the judgment of the supervisor in charge of the examination will be final and without appeal." Document No. 2 does not mention the possibility that an irregularity might be discovered after the test administration; anticipated or not, such troublesome events soon occurred.

Some indication of the extent of cheating discovered on early College Board tests is provided by an archival gem: a stenographer's transcript of a conference of supervisors called together in 1926 to discuss the forthcoming

first administration of the Scholastic Aptitude Test. Commenting on the candidate misconduct problem, the Board's Secretary, Thomas Fiske, stated, "I do not believe that the half dozen cases that we discover [every year] are an insignificant proportion of the real number of cases that exist."[2] A few minutes later, Fiske commented:

> In Boston and Cambridge... there have been... a good many cases of impersonation and cheating. One year... two boys were expelled from the Boston Latin School... and their parents raised a terrible row and said, "Why, you are punishing our boys, ruining their careers, and all they are guilty of is the thing on the basis of which Mayor Curley was elected Mayor of Boston." Curley was in jail for impersonating other people at Civil Service examinations, and when his political enemies attacked him, he patted himself on the chest and said, "Why, I think a fellow who would go to jail to help another fellow get a job must be a pretty good sort of fellow," and the people of Boston agreed....[3]

While some attempted impersonations, other test takers brought crib sheets into the testing room. Fiske seemed amused that every year readers found reference materials in the answer books of forgetful candidates. In 1925, one student explained that, at the suggestion of his teacher, he had removed pages from his geometry textbook in order to study them on the streetcar he took to the examination center. During the test, when he removed his handkerchief from his pocket, the pages, he assumed, must have fallen unnoticed into his answer book.[4]

Fiske described the Board's policies and procedures for dealing with these occasional cases as follows:

> In cases of suspected cheating, we always try to get the advice of the supervisor and the head of the preparatory school from which the candidate comes. Whenever we suspect a candidate of having been guilty of a violation of the rules, we always investigate as fully as possible and consult everybody who could possibly shed light on the subject....

> We have a rogue's gallery [in the office], but it does not really mean that the boy is a rogue. It means he is suspected of an irregularity or a violation of our rules. We have to refrain from actually accusing candidates of dishonesty....[5]

Because most College Board candidates of that era were preparatory school students well known to the test supervisors, impersonations must have been quite unusual. But Fiske probably underestimated the extent of other forms of cheating. For every candidate who left a crib sheet in his answer book, there must have been many others sufficiently composed to take their textbook excisions home.

Fiske's willingness to discuss investigations in progress with school officials differs strikingly from current practice, which reflects a greater concern for test taker privacy. More consistent with current policy is Fiske's distinction between dishonesty and rule breaking. As will be seen below, ETS has continued to avoid characterizing candidate misbehavior in legal terminology.

It seems curious today that Fiske did not specify copying from other papers as a cheating method. At the time, Fiske had little experience with the "new-type" multiple-choice tests; in the absence of effective deterrents, short answer tests facilitate illicit transcription. The multiple-choice test format created a potential test security problem soon recognized by specialists.

In 1927, a year after the supervisors' meeting, Charles Bird of the University of Minnesota published his article, "The Detection of Cheating in Objective Examinations," perhaps the earliest explanation in print of the basic rationale still used for making determinations in copying cases: "We can tell whether the identical wrong answers in two papers exceed a number which is possible by chance." Although his statistical procedures were crude

by modern standards, current methods share Bird's assumption that "a student who secures information surreptitiously from another paper is seldom capable of discriminating right from wrong answers in an objective examination."[6]

Despite the availability of this technique, it does not appear that the College Board immediately felt in need of it. After the introduction of the Scholastic Aptitude Test, the Board's written procedures for dealing with cheating continued to be limited to admonishing candidates to behave themselves and encouraging invigilators to be on the qui vive. Remarkably, in the first decade of SAT administrations, not one candidate's score was cancelled for misconduct. In fact, according to Cecil Brglyer, a central figure in SAT affairs from 1927 to 1936, there were not even any investigations.[7]

For a number of years after ETS was established in 1947, testing program directors were responsible, with the assistance of test administration staff, for investigating questioned scores and determining whether to invalidate them. What evidence there is suggests that investigations were quite unusual.

In 1955, a Task Force on Physical Security of Tests, consisting of William Bretnall, Catherine Sharp, Ned Terral, and William Van Cleve, made an in-depth survey of test security at ETS and recommended that a Security Officer be appointed.[8] Mr. Van Cleve, who worked in Supervisor Relations, was named to the new position in 1956. Although the task force report had not mentioned impersonations or communication (its concern was to minimize test book losses), he was given responsibility for handling all test security problems. During his first year, there was a difficult impersonation case in which Bretnall's father, a private detective, lent a hand. As his son recalled:

> The case began in August of 1956 and dragged on through the end of January 1957. In between were continuing threats of a law suit, appeals to Governor Ribicoff of

> Connecticut, accusations that we were persecuting Armenians in general, interviews by my father with 25 or 30 people, consultations with lawyers, and an endless stream of letters, memoranda and phone calls. Finally, six months after it all began, the candidate and his impersonator admitted their deed. It was a great relief to all concerned.[9]

Such experiences indicated that ETS needed a trained investigator as Security Officer. Perhaps this change, made in 1958, was responsible for a dramatic increase in the number of confirmed impersonations: from nine in 1957-1958 to fifty-nine in 1958-1959. In the latter year, another five scores were cancelled for cheating by other means.[10]

In 1966, the ETS Security Officer, Captain Paul D. Williams, a former Navy officer with intelligence experience, described the procedures followed by the Test Security Office.[11] Williams stated that after gathering the available evidence, including an opinion from a handwriting expert if impersonation was suspected, he would contact the candidate, usually through a school official. (Until about 1960, the interview was almost always conducted in person, but by 1964-1965, the case load had increased to the point that 95 percent of the discussions were held over the telephone.) Then, as Williams continued:

> The candidate is apprized of the investigator's identity, the reason for the investigation, a resume of the evidence indicating that the candidate has disqualified himself; namely the score comparison, the indication of impersonation, or of copying, as the case may be, and any other pertinent observations the investigator has made. The candidate is then requested to comment and if he confirms the purport of the evidence, a statement affirming the irregularity and agreeing to the invalidation of the questioned scores is dictated to the candidate. He is requested to forward this statement in writing to the Security Office....
>
> If he should refuse to confirm the evidence and indicate positively that he had not been involved in any irregularity, the investigator will suggest to him that he retake the test at a special administration.... If

> the candidate refuses the opportunity to be retested, the investigator will inform him that a complete account of the investigation will be made to each of the institutions which have received his questioned scores.

If the candidate would not agree to cancel his or her score and refused to retest, ETS would send the score report to the score recipients with a "dubious validity letter" which "would draw no conclusions," although it would describe the evidence suggestive of misconduct. Use of the dubious validity letter was infrequent since most confronted examinees agreed to cancel their scores: 85 percent in 1964-1965.

Some score recipients, notably law schools, wanted ETS to do more. In 1964, the Law School Admission Council's Executive Committee appointed an Ad Hoc Committee on Test Security to look into problems of cheating and to review ETS investigative practices. Indirectly, this Committee contributed to the formation of the ETS Board of Review.

As compared to test sponsors of other ETS programs in the 1960's, the Law School Admission Council had already shown a significant interest in infractions. For example, the LSAT was the only program which required Test Security to check the validity of increased repeater scores as a standard procedure; at this time a gain of 100 points or more triggered an investigation.[12] The law schools' stated reason for their particular concern was that they were required to report on the character of their graduates when they applied for bar examinations.

The LSAC Ad Hoc Committee presented its findings to the Council in 1965. Their report summarized existing ETS procedures, including the use of comparisons of wrong answers in copying cases, and termed them "adequate." The Committee did make a number of recommendations for change, most of them minor,

but of particular interest was a proposal that if a candidate persisted in denying misconduct, ETS should arrange "a hearing panel of three persons through the American Academy of Arbitrators or the local bar association" to make a final decision.[13]

In response to the Committee's report, ETS Counsel John Graham accepted some of the suggestions, including sending "an information copy to the testee of any final letters to score recipients about the validity of his score." But Graham believed that some of the Committee's other suggestions would make the investigations inadvisably resemble criminal proceedings. "ETS," stated Graham, "has formulated its investigative procedures to provide the maximum protection to the integrity of reported test scores consistent with basic fairness to the testee and minimization of potential liability against ETS... every effort is made to avoid the implication that the investigation is a criminal proceeding. For example, statements obtained from suspected testees are called admissions, not confessions."

Graham "strongly opposed" the Committee's suggestion of a review board, not so much over the score validity vs. misconduct emphasis, but for what he believed were practical reasons: the expense; the administrative burden and psychological effect on test security personnel whose actions would be subject to review; the abuse of the procedure, such as for a delaying tactic; and the likelihood that the board's decision would not be binding and provide no protection against litigation.[14]

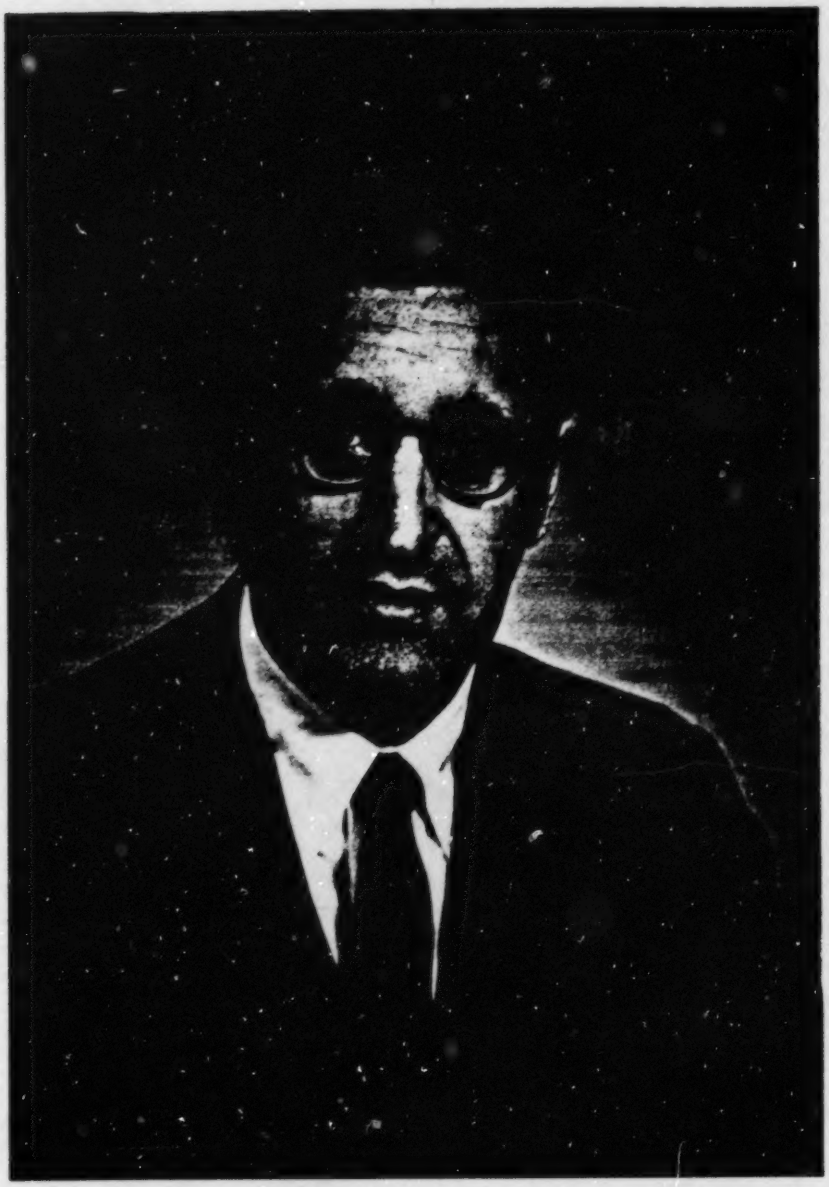
As a result of Graham's comments, the LSAC Executive Committee requested the Security Committee to reconsider its report and consult further with ETS. In 1966, the Security Committee reported back with revised proposals that represented a compromise between its initial recommendations and Graham's

objections. But the Committee stood firm in its desire for a hearing board for disputed cases, as did Graham in his opposition.[15]

The law school concerns led ETS Vice President Robert J. Solomon to ask Robert E. Smith, Executive Director of General Programs, to take a close look at the test security operation. Subsequently, Smith and Solomon agreed that changes were needed. As Smith recalled, "One thing that bothered us in the very beginning was that, in our view, there was no opportunity for the youngster really to get good advice" before responding to ETS's interrogator. "We were also concerned," continued Smith, "about the business of telling the test takers that the thing they needed to do" was to sign a dictated confession. Nor did they like the dubious validity letter. Even though the existing procedures were based on a concern focusing on score validity rather than wrongdoing, they felt the result was overly punitive and, as Smith concluded, "We should refrain from any actions that made the business more public than was absolutely necessary."[16] In order to develop viable alternatives, Solomon appointed a Committee to Review ETS Security Procedures, chaired by Smith, in November 1966.

In his "Presentation on Cheating" to the ETS Board of Trustees on May 7, 1968, Smith presented his committee's two alternatives. Smith's preferred possibility was to advise score recipients when a question arose concerning a candidate's score, before an investigation which would result in a determination of the score's validity. If the institution wished, it could contact the candidate and ask for a retest. ETS would not have to cancel any scores.

The Trustees strongly rejected this radical proposal,[17] stating that it was ETS's responsibility to determine if scores were valid, and asked ETS to proceed along the lines of Smith's other alternative, which featured some of



Bob Smith, 1970

the procedures introduced later when the Board of Review was established: no request that candidates sign a statement of admission; offering all candidates an opportunity to retest; refunding fees to those whose scores were cancelled; and providing no explanation to score recipients of the reason for a cancellation. However, in this proposal, final determinations of score validity would be made jointly by the testing program director and the Security Officer. "A hearing board could be resorted to where unusual circumstances recommended it."

By August 5, 1968, Smith and others at ETS had determined that it would be preferable to have an internal board, independent of both Test Security and Program Direction, make final decisions in all cases, leaving other features of this plan largely the same.[18] In December, the Trustees approved this plan, codified as the "General Policy on Questioned Test Scores" and the ETS officers appointed the Board of Review. Charter members were: William A. Angoff, Marion G. Epstein, John S. Kramer (ETS Counsel), and Robert E. Smith, Chairman. The first formal meeting of the Board was held on January 9, 1969.

The policy states, that in order for ETS to fulfill its "obligation to deal fairly with the candidate as well as to assure the authenticity of the scores to the recipients," the available evidence concerning a questioned score is presented [by the Test Security Office] to a "Board of Review consisting of the ETS legal counsel and three senior professional staff members not directly responsible for the administration of the test programs concerned." Three members of the Board constitute a quorum. If there is unanimous agreement that the validity of a score is in doubt, a registered letter is sent to the candidate, stating that the Board will cancel the score unless the candidate can confirm by a retest or an adequate explanation.

At the heart of the policy is the fundamental principle that "the proper interest of ETS rests with the authenticity of the scores it reports and not with providing evidence or a judgment of misconduct by candidates.... Punitive intent toward the candidates" should be avoided. Accordingly, even if the Board of Review has strong evidence of test taker misconduct, it confirms a questioned score if the individual can demonstrate that the score is valid; in fact, it will not question a probable miscreant's score in the first place if available evidence suggests that retesting would confirm. If scores are cancelled, the candidate's fees are returned and the scores removed from ETS records. Score recipients are not provided the reason for the Board's cancellation of the score.[19]

Most specifics of the General Policy are still in effect. Differences concern the size and composition of the Board (more members almost immediately and the exclusion of legal counsel in 1980) and the following supplemental features.

The first major addition to the policy occurred in the Board's second year. For some time, Test Security had been using the computer to identify the LSAT candidates with score gains sufficiently large to warrant further investigation. At their Spring 1968 meeting, the ETS Trustees had suggested that ETS investigate the feasibility of expanding the application of this technique to other programs. In February 1970, the Board of Review adopted its "Recommended Supplementary Policy on Questioned Test Scores," which mandated score gain checks by computer in other national programs. The Scholastic Aptitude Test was the first new program to which the technique was applied. The triggering cutting (difference) score for the SAT was established by research conducted by Mr. Angoff, who had found that above a certain score

gain on the December 1968 SAT, 88 percent of the answer sheets demonstrated clear evidence of impersonation or communication.

The Board noted in its new policy statement that a major advantage of computer-identified cases was that it permitted scores to be questioned before they were reported; the result was less embarrassment for those whose scores were subsequently cancelled. Another effect, of course, was to identify cases which might never have been discovered. Significantly, the Board immediately adopted a rule never to cancel a candidate's score on the sole basis of an improbable score gain; it directed Test Security not to present such cases to the Board if no other evidence suggestive of invalidity could be found.[20]

Despite the Board's attempts to be fair to those whose scores were questioned, some candidates and parents responded quite negatively to the Board's initial letter. To ameliorate this problem, the Board tried a short-lived "pre-investigation procedure," as described in a policy statement Smith issued in February 1975.[21]

Under the new procedure, Admissions Testing Program and Graduate Record Examinations candidates whose scores were flagged due to large score differences were sent a letter informing them that an investigation would soon be initiated. The letter also offered an opportunity to cancel or retest before an initial judgment was made concerning the validity of the score. The supposed advantage of this procedure for the candidates was to "minimize any accusation of misconduct" and for ETS, to save investigative costs for those who authorized cancellation. Despite careful wording, some students perceived the letter as an accusation and felt that ETS should investigate further before contacting them; reaction was sufficiently negative that the procedure was dropped after a few years.

In 1976, as a result of a review of ETS test security practices by the College Board, questioned scores which had not been reported began to be placed in a "suspense" file until, if ever, the test taker responded in writing to the Board's initial decision to question the score. In 1984, the Board of Review returned to its earlier policy of cancelling scores if the candidate does not respond, but with a longer waiting period when scores have not been reported.

A more lasting innovation officially introduced in 1978 is to offer the opportunity to have all of the evidence sent to the institution designated to receive the score, if the institution agrees.[22]

Another innovation, still current, is arbitration, proposed initially, it may be recalled, by the LSAC Ad Hoc Committee on Test Security in 1965. For LSAT candidates, the option of submitting the case to an arbitrator appointed by the American Arbitration Association was made available in 1975. In 1981, the College Board agreed to make arbitration an option in its testing programs and the procedure was adopted subsequently in the GRE and GMAT programs.

In addition to these policy changes, there have been refinements to internal Board of Review procedures and, under its direction, those of the Test Security Office. Perhaps most significant has been the development, application, and increased reliance on statistical indices used to provide a probability estimate that the correspondence of incorrect responses in a pair of answer sheets could have occurred by chance. The first such indices were developed by William Angoff in 1970 and revised in 1973 by Louis Lavine, who in 1976 became the second chairman of the Board. In 1979, Frederick Kling developed the currently used Index K. Recently, Index K has made feasible a serendipitous method of identifying cases to be brought to the Board for a

decision. These "developed cases" are discovered occasionally when the computer is used to run comparisons of all possible pairs of a group of answer sheets to determine the probable source, if any, for a test taker suspected of copying. (The all-pairs technique is also used when collusion or pre-knowledge of a test is suspected of an indeterminate number of candidates.) If strong K's are found, Test Security submits the developed cases to the Board.

The use of increasingly sophisticated statistical methods and computer technology has had several effects on the Board's work. First, although some scores continue to be questioned by score recipients, the majority of suspected irregularities are discovered first at ETS and are resolved before scores are reported. Second, the computer has significantly increased the number of cases coming to the Board, despite additional measures implemented to prevent irregularities at the test centers. This increase to about 2,000 cases annually has necessitated a gradual expansion of Board membership to seventeen; the growth is in large measure a consequence of the computerized score difference check, although increased candidate volume and other less tangible factors may also be involved. Third, the score difference check and Index K may have figured in an increase of the proportion of cases involving probable communication to almost half, with the majority still impersonations.

The nature of the changes to Board procedures suggest that although the basics of Board policy have remained unchanged, there is a continuing effort by both ETS and test sponsors to review and refine Board practices. The courts have also provided evaluations of ETS test security policies and have found them satisfactory, even laudable. Although threats of lawsuits have been frequent, only a few cases have come to trial. In one early case before the establishment of the Board, involving probable preknowledge of the National

Teacher Examinations, the judge concluded, "The evidence is circumstantial, but circumstantial evidence is not to be excluded. Circumstantial evidence is a basis for findings much more severe and much more harsh than those that would result in this case. To deny circumstantial evidence is to deny the ability of the human mind to reason."[23]

Court cases after the Board's founding have also upheld ETS policies. In 1983, there was a considerable amount of public attention concerning a probable preknowledge case involving four SAT candidates. A key issue was whether ETS had to prove that cheating occurred in order to have a reasonable basis to question the candidates' scores. Finding for ETS, the judge concluded that "ETS may not properly be forced to inquire into questioned scores in the manner of a law enforcement agency or to adjudicate guilt or innocence in the manner of a court."[24] This decision, upheld on appeal, further confirmed the wisdom of the key element in ETS' test security policies, developed in the 1950's and enhanced by the establishment of the Board of Review in 1969, that ETS' appropriate concern in this area should be score validity. The Board continues to function on this principle.

Gary D. Saretzky
September 7, 1984

Footnotes

[1]Misconduct on examinations is undoubtedly as old as testing itself but its emergence as a potentially large-scale problem probably occurred when standardized tests were introduced in China as a requirement for government service. Particularly in their mature form during the Ming to Qing dynasties (1368-1911), "the examinations were in fact a cornerstone of the social and political edifice... and the centrepiece of the life and lore of the scholar-gentry-official class" to which success on the tests provided entry. Recognizing that the temptation to cheat was very great, the authorities instituted extraordinary preventive measures. For the provincial and metropolitan examinations, test administrations were conducted in huge walled compounds resembling prisons, tightly supervised by armed guards and examiners. Candidates were stripped and searched before admittance, then locked up in virtually bare cells for up to four days of testing. There, each morning, they were brought their food and examination questions by deaf mutes, who would pick up the papers at the end of the day.

Despite these stringent prophylactic policies, some examinees attempted to achieve higher scores than they deserved. Occasionally, test development officials (and their unlucky subordinates) were beheaded for accepting money to reveal questions in advance. Another technique was to bribe the deaf mute either to allow reference materials to be brought in (baked in cakes, for example) or to take questions out to be completed by a confederate in an adjoining cell and returned. Impersonation was also practiced, as were underhanded methods to influence the grading process. (Bernard Luk, "The Civil Service Examinations in Late Imperial China," Orientations, 13, 21.)

[2]"Conference of Supervisors Held in the Trustees' Room, Columbia Library, Saturday, January 9, 1926, at 10:30 a.m.," Benjamin Gotthelf, Shorthand Reporting, 154 Nassau Street, New York City, 93.

[3]Ibid., 103-104.

[4]Ibid., 92.

[5]Ibid., 94, 98.

[6]Charles Bird, "The Detection of Cheating in Objective Examinations," School and Society, February 26, 1927, 261-262. Robert E. Smith, the first chairman of the Board of Review (1969-1976) studied under Bird in 1949.

[7]Based on review of annual reports on the SAT in the published annual reports of the College Board 1927-1935 and personal communication from Cecil Brolyer, July 2, 1984. Brolyer assured me that the occasional special score reports issued because candidates "misread the directions" were not due to cheating.

[8]"Report of the Task Force on Physical Security of Tests," Draft, [July 1955?], in Henry Chauncey Papers, Folder 388.

[9]William B. Bretnall, "An Aspect of Test Security, 2200 B.C. to 1966 A.D.," Talk at the College Board Staff Meeting, Skytop, Pennsylvania, June 2, 1966, 4.

[10]Based on annual reports of the Test Security Office.

[11]Memorandum, "Statement of Existing Procedures for Disposition of Discrepant Scores in all Testing Programs," July 25, 1966.

[12]"Report of the Ad Hoc Committee on Test Security," June 1965, in LSAT Papers, I.11.53.

[13]Ibid.

[14]"Comments by Educational Testing Service on Report Recommendations of the Law School Admission Test Council Ad Hoc Committee on Test Security," presented December 3, 1965, in LSAT Papers, I.11.53.

[15]"Second Report of the Ad Hoc Committee on Test Security," in LSAT Papers, I.11.53.

[16]Robert E. Smith Oral History, February 24, 1984.

[17]Ibid.

[18]A memorandum from Smith to Solomon on August 5, 1968, states, "I would like to have the Board of Review appointed...." This is the earliest reference I found to the Board of Review. Attached to this memorandum was a draft policy very similar to the one approved by the Trustees in December 1968. It should be noted that the LSAC continued to require that an irregularity report be sent to the designated law schools along with cancellation notices. This program also had several other unique test security practices. Several years ago, the LSAC began conducting its own test security investigations and LSAT cases no longer come to the ETS Board of Review.

[19]Cancellations continue to be made for misconduct alone or irregularities unrelated to candidate behavior; the Board of Review is not responsible for such cancellations, the reason for which may be reported to score recipients if required by the test sponsor.

[20]Arthur L. Benson, "How ETS Deals with Test Scores of Questionable Validity," October 8, 1971, 4.

[21]Robert E. Smith Oral History, op. cit.; Robert E. Smith, "ETS Policy and Procedures: Test Scores of Questionable Validity for Candidates in National Testing Programs," February 1975. Smith believes that the pre-investigation procedure was responsive to parents who stated that they would have preferred to be given the option of cancellation earlier in the process before any investigation, simply on the basis of a machine-identified score gain.

[22]"ETS Procedures for Determining the Validity of Questioned Scores," March 1978, 6. This procedure was used earlier at the discretion of the Board of Review.

[23]George C. Young, Civil Action Opinion, U.S. District Court, Southern District of Florida, Miami Division, Docket Number 63-449-CIV-EC, Decided May 21, 1964, 13.

[24]Richard S. Cohen, Civil Action Opinion, Superior Court of New Jersey, Chancery Division, Middlesex County, Docket Number C-1715-83, Decided August 4, 1983, 47.